Independent Peer Review

Benchmark stock assessments of lingcod and Pacific Ocean perch

Prepared by:
Panayiota Apostolaki

Assignment undertaken for:
Center for Independent Experts, CIE, US

August 2017

Table of Content

Executive Summary

The panel review meeting for the benchmark stock assessments of lingcod and Pacific Ocean perch (POP) took place in Seattle between June 26th and June 30th, 2017. The focus of the review was the stock assessment work done for the two stocks of lingcod and one stock of POP.

Both species were assessed using the length- and age-structured modelling software called Stock Synthesis, which has been extensively used for stock assessments in the West coast of the US. The assessments made use of a diverse set of data sources to capture the best scientific knowledge for the two species, including fishery dependent and fishery independent abundance indices, and length and age composition series. There was also further discussion at the review meeting about the best parametrisation of the models that led to the STAT proposing different model parametrisations from the ones put forward at the beginning of the review meeting.

Many sources of uncertainty were captured in the analyses and the results also highlighted the fact that there was limited information in the input data and/or they were supporting differing status trends. The updated models reflect the data and processes in which the STAT team had more confidence, and although a number of parameters are still not well defined, sensitivity analyses provided a useful perspective into the effect they have on model results.

The outcomes of the final base case model for POP suggested that the stock is at pre-exploitation levels while for lingcod, the estimate of the size of the stock was smaller than that found in the previous stock assessment for both the Southern and Northern stocks. However, the Northern stock is predicted to have recovered. The Southern stock is below 40% of its pre-exploited size.

Overall, the assessments represent commendable effort to provide the best assessment frameworks for the two species considered, and make use of the best scientific information.

Background

The 2017 Benchmark stock assessments for lingcod and Pacific ocean perch (POP) focused on the stocks of lingcod and POP that are found along the west coast of the United States from California to the Canadian borders.

Benchmark stock assessments have been carried out in the past for both species; the last benchmark for lingcod took place in 2009 and for POP in 2011. For the latter, previous assessments indicated that the population was depleted and as such, POP has been managed under a rebuilding plan for over a decade. A rebuilding plan for Lingcod was also in place between 2000 and 2006, but after that the population was considered to be rebuilt.

Both species were assessed using the length- and age-structured modelling software called Stock Synthesis (SS). The software has been extensively used for stock assessments in the West coast of the US and elsewhere, and aims to provide a framework for combining information from different type of data to inform the model results about the status of the stock and impact of fishing pressure. The software includes two components; a population dynamics sub-model that simulates the age and length-specific structure of the population and an observation sub-model which can make use of a wide range of data to calibrate the model. The observations that can be used in SS include: fishery CPUE or effort; survey abundance; discards; length-, age- and weight-composition data; and tag-recapture data (Methot and Wetzel, 2013)

A number of fisheries specific and non-fisheries data series were employed in the assessments and the models simulated fishing pressure associated with both commercial and recreational fisheries. For lingcod, commercial landings were split into two categories; one coming from vessels using primarily trawl gear, but also includes other net gear that had a much smaller contribution to catches; and one from vessels using fixed gear such as longline, troll, and hook and line. The model used three fishing fleets to simulate commercial pressure for POP; a combined bottom trawl, mid-water trawl, and fixed gear, a historical foreign fleet, and the at-sea hake fishery.

Reconstruction of catch data also took place prior to the stock assessments and that produced a long catch series covering about 100 years for both species. Fishery dependent and fishery-independent CPUE series were also used in the stock assessment covering a shorter period of time than that for the catches; for lingcod, CPUEs covered the recent period going back to about 30 years while for POP, one commercial CPUE series covered earlier years (prior to 1980) while there were 5 CPUE series covering the period after 1980s. The assessments also utilised length and age composition data that were available for both species.

The assessments made use of available data on growth, maturation, and fecundity to produce equations to simulate those processes. However, information about some elements of the species biology, such as natural mortality, was limited and that was one of the sources of uncertainty.

The assessment calculated reference points based on SB40%, SPR 50%, and MSY that reflected targets used for the management of these fisheries. The assessment also did sensitivity analysis to test the effects of uncertainty on model results as important parameters could not be estimated and had to be fixed. Retrospective analyses and projections under different combinations of values of selected model parameters and for different future catch quotas were also run. The latter produced decision tables with estimates of ABC under different states of nature.

A number of changes were made to the original model to correct errors, modify assumptions about some of the processes (e.g. recruitment deviations), and update the type of input data used (e.g. exclusion of some CPUE series). The final proposed models represented the version of the model in which the STAT had more confidence.

For POP, the model results indicated that the population has recovered to almost pre-exploitation levels and projections suggested that it might go even above the average pre-exploitation stock size. Concerns were expressed about the realism of the results, and sensitivity analyses were part of the process to ensure that all plausible states of nature had been represented. Overall, the stock seems to be at a better state than predicted in the previous assessment and its recovery seems to be supported by good recruitment.

For lingcod, the population size estimated for both stocks was smaller than that found in the previous stock assessment, but the results were more positive for the Northern stock that is estimated to be above 50% of its pre-exploited size. The Southern stock appears to be more depleted and below 40% of its pre-exploited size.

Two CIE reviewers were commissioned to participate in the stock assessment review panel and conduct an impartial and independent peer review of the stock assessments of the two species, and in accordance with the SoW and ToRs herein. One of the reviewers also acted as the "consistent" CIE reviewer and participated in all STAR panels held in 2017. Each CIE reviewer is also required to produce an independent peer review report in the format and content of which is described in Annex 1. The report should be addressing each ToR as described in Annex 2.

I was the consistent reviewer, and this document provides my review of the 2017 benchmark stock assessments of lingcod and Pacific ocean perch. Further details on the reviewer's role and the review request of the Center for Independent Experts are presented below and in Appendix 2.

Description of the Reviewer's Role in the Review Activities

I was contracted to:

1) Conduct necessary pre-review preparations, including the review of background material and reports provided in advance of the peer review.
2) Participate during the STAR Panel 1 review meeting scheduled in Seattle, Washington during the dates of June 26-30, 2017 as specified herein, and conduct an independent peer review in accordance with the ToRs (Annex 2)
3) No later than July 14, 2017, submit the draft independent peer review report to the contractor. The CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in Annex 2. (Appendix 2).

In addition to that, in my role as an active and engaged participant, I voiced concerns, suggestions, and improvements throughout the panel discussions while respectfully interacting with other review panel members, advisors, and stock assessment technical teams.

Summary of Findings

TOR 1. Become familiar with the draft stock assessment documents, data inputs, and analytical models along with other pertinent information (e.g. previous assessments and STAR panel report when available) prior to review panel meeting.

Several documents were provided two weeks before the meeting for both species including:

- The draft stock assessment reports;
- The Pacific Fishery Management Council's Scientific and Statistical Committee's Terms of Reference for Stock Assessments and STAR Panel Reviews;
- Stock Synthesis (SS) Documentation; and
- Past STAR Panels report and individual CIE reviewers' reports.

The full set of bibliography that became available to us either before or during the meeting is described in Appendix 1.

I reviewed the assessment reports prior to the STAR Panel meeting and became familiar with other documents provided including the analytical model (Stock Synthesis) and the data that were used to populate the model. That process highlighted a number of questions which formed part of my contribution to the meeting.

TOR 2. Discuss the technical merits and deficiencies of the input data and analytical methods during the open review panel meeting.

Lingcod

The assessment of Lingcod assumed that there were two stocks; one that occupies waters off Washington and Oregon (northern stock) and the second covering waters off California (Southern stock). No mixing of the stocks was simulated.

Catch and CPUEs
The data used for the 2017 assessment included catches from commercial fisheries (trawl and fix gears modelled separately) and recreational catches for both the Northern and Southern stocks. Catch reconstruction led to expansion of catch series to before 1900 for the Northern stock and 1930 for the Southern. Length and age composition data were available for all fisheries, both recreational and commercial. Data on discards from the fix gear and trawl fishery were also available for recent years (after 2000) and were included in the calculations. During the meeting, it was confirmed that selectivity that could change over the years was also included in the model and different options for blocking were discussed leading to some changes of the blocks used.

At the beginning of the review, the model included the following CPUE indices:

- AFSC Triennial Trawl Survey Index, this survey was split into two components, an early one and a late one. Both provided length compositions but age composition was only available for the late part (after 2000). Also, concerns were raised about the results of the survey in 2004 as the approach followed was different from that used in previous years.
- NWFSC Trawl Survey Index, this survey also produced length compositions for both the Southern and Northern stocks.

- NWFSC Hook and Line Survey Index, also produced length composition data.
- PacFIN Trawl Logbook index, this also provided age and length composition data.
- OR Nearshore Commercial Fixed Gear Index, including age and length composition data.
- WA Dockside Recreational Index, including age and length composition data.
- OR Dockside Recreational Index, this study has many more samples than the OR Onboard Recreational but there is less confidence in the data due to recording issues.
- OR Onboard Recreational Index.
- CA Onboard Recreational Index.
- Central CA Dockside Recreational Index – this index was used only in a sensitivity run.
- Lam research composition data provided length and age composition for 2 years. The data from this series were inputted in the model as conditional age at length for both the Southern and the Northern stock.

All the recreational CPUE indices and the Oregon fixed gear commercial fishery index were standardized using a standard delta GLMM while all fishery independent (surveys) indices and one commercial index, PacFin trawl logbook index, were standardized using a geospatial delta GLMM package called VAST. VAST has been used by the STAT in the past and, I understand, has also been presented at the SSC. The methodology has also been published (Thorson et al. 2015), and it appeared to be the preferred standardisation tool for the Northwest fisheries science center (NWFSC).

Data from the two West coast-wide surveys (AFSC Triennial and NWFSC Trawl) were also used in a single standardization run which produced two separate indices, one for South and one for North. This was questioned (see ToR 7) and additional runs separating the data for North and South were also done.

Further, there were some unusual patterns in the age composition data of the WA recreational survey suggesting that male fish are bigger than female fish. Following the review meeting, further work revealed that there is high chance that sex specific data have been incorrectly reported during the data collection process but furthermore, the gender of the individuals in the dataset were incorrectly inputted into the model. It is not clear what impact the latter issue will have in terms of how and whether those data should be used in the baseline model (these data have been excluded from the baseline at the moment) but, at the time of writing this report, the STAT had indicated that a sensitivity analysis with the corrected set of data would be added to the final version of the assessment report.

Biological information

Age information was available from studies that used spines to age lingcod as reading otoliths was not considered a reliable method. However, even with the spines, there are some concerns about inter-lab error as two labs have done aging for this species (only recent years though). Similarly, there were concerns that the way sex allocation of fish from fisheries data was done for unsexed fish was incorrect and could lead to age distributions that did not represent the selectivity pattern of that gear. Analyses taken indicated that some bias towards younger fish might exist so, that is a source of uncertainty. There is also uncertainty about the spawning behavior of lingcod; for example, it is not clear whether females produce multiple batches of eggs and/or

contribute to separate nests. With regards to the two stocks, the assumption is that the stocks do not mix, but it was indicated during the meeting that there are currents that could transfer larvae from one area to the other.

A prior for natural mortality as well as steepness were created based on analyses that was not used in the previous assessment so, is there another addition to the model although the final parametrisation used fixed values. Length at 50% maturity was also updated from the previous assessment and relied on data collected in the past four years. The 2017 length at 50% maturity is lower than that used in 2009 for both stocks. It is of concern that the new calculations only used data from recent years after the population had been exposed to considerable exploitation.

Pacific Ocean Perch

Catch and CPUE series
Catch series from three fisheries representing fisheries that target POP (both by US and, in the past, by foreign vessels) and the at-sea hake fishery and data on discards were used to capture removals. On discards, two sources were utilised; one covering discards between 1985 and 1987 and a more recent one that includes discard data starting in 2003. The data from the direct commercial fishery were used to develop a CPUE series as well as length and age composition series. Length and age composition data were also available for the at-sea hake fishery.

At the beginning of the review meeting, the proposed model included the following information to provide signals about the stock structure and depletion:

- Triennial shelf survey that provided a relative abundance index, length composition and age composition data for the period pre-2005;
- A cooperative research survey of POP off Washington and Oregon which covered only two years (1979 and 1985);
- Two slope surveys; one of which provided a CPUE index and a small number of length composition data (AFSC) and the other provided a short series of CPUE and only two years of age and length composition data (NWFSC slope survey);
- A slope-shelf survey (NWFSC shelf-slope survey) which is the extension of the NWFSC slope survey and provide a CPUE series for recent years as well as length and age composition data;
- A historical CPUE series from the domestic fishery from the INPFC Vancouver and Columbia areas;
- A data series of catch and length compositions from commercial fisheries extracted from PacFin and which provided samples from Oregon and Washington but not from California.

All the fishery-independent CPUEs were standardised using a spatio-temporal delta-model that was implemented in an R package called VAST.
The discussion during the review indicated that some misreporting might have affected the quality of catch data; in particular, there are concerns that California catches that were landed in Oregon might have not been captured correctly in the official recording system.
Also, the STAT reported that catches from the shrimp fishery were not included in the model as they were so small. However, as that fishery takes small fish, its impact might be more than the catch in biomass suggest. Additional calculations showed that removals from this fishery could be as much as 1% of the number of young fish. The number of fish caught seems to go up when a good recruitment cohort is present and

therefore, it is worth monitoring the catches and take their impact into account especially when a large new cohort has been detected.

Biological information
Otoliths were used to calculate maximum age but only from a subset of the otoliths that were read using the burn and break method as scales and surface reading methods were considered to be unreliable. That elimination process aimed to reduce bias from reading but also reduced the sample size.

The maximum age was an important element of these analyses as it was used to calculate natural mortality. In the first model presented to the reviewers, the age of 100 years was used to calculate mortality. The final model adopted assumes max age of 60 years highlighting the uncertainty that there is about maximum age. This is also supported by comments during the meeting that indicated that there had been fish even older than 100 years that have been recorded, but there is doubt about the quality of those high estimates of age. The overall perception is that the uncertainty in age readings increases very quickly as we move to fish older than 30 years.

A sex ratio of 1:1 was assumed for the calculations and both maturity and fecundity estimates were updated based on data that became available since the previous assessment. Growth was gender-specific and was calculated using fishery dependent and fishery independent data. However, the analyses then used data inside the model to produce model-estimated growth curves that reflected the information included in the input data. Conditional age at length data were used to support the process and that assumed that the age at length data represented a random sample of the population and thus, it represented different ages equally. However, analyses conducted during the meeting suggested that sampling was not random, at least in some cases, and therefore, might be misrepresenting the structure of the population.

A similar concern was raised about the North-South bias as a lot of the data for the model came from the Washington and Oregon area and it was not clear to what extend they represented the biology and population structure in the South (e.g. no age structure data from the fishery in the South). This was an issue that could not be resolved during the meeting as further work will be needed to address/evaluate this.

TOR 3. Evaluate model assumptions, estimates, and major sources of uncertainty.

Both assessments used the length- and age-structured modelling software Stock Synthesis (SS) but for POP, it was assumed that the population found in the west coast of the US constitutes a single stock, while for Lingcod the assumption was that there were two stocks. The models for both species simulated a sex-disaggregated population dynamic. The models made use of a diverse set of data covering landings and discards, CPUEs, length- or age-composition data, and length specific maturity and fecundity. They also used the Beverton-Holt stock-recruitment relationship to link spawning potential to recruits and selectivity, and retention functions and blocking to characterise the behaviour of the fisheries over the years.

Lingcod

Models for both stocks were run from pre-1900s onwards but only the Northern stock had catch data that extended that far back. The Southern model was run starting from such an early year only because that feature was already there from the Northern model. The model that was proposed during the first day of the meeting assumed fixed

growth for the Northern stock and used the values for M and h that were estimated for the Northern stock as input to the Southern stock. This was because the Southern model did not have enough information to estimate them. However, the latter assumption has little basis, especially given the original assumption that the Southern stock is completely different from the Northern stock.

Different parametrisations were also considered for the Northern stock as the model could not estimate steepness, natural mortality and growth at the same time. For the Northern model proposed originally, growth was fixed. Also, the model estimated recruitment deviations only for the recent years. However, the final model selected as a base case included deviation in recruitment from early years, and also fixed the values for natural mortality and steepness.

Some unexpected large males in early years in sex-specific composition data from PacFin also seemed to be the results of model assumptions; in particular, assumptions in the size composition package used for PacFin which allocated unsexed fish 50:50 into males and females. That choice is not supported by differences in biology and dynamics of males and females, and probably does not account for upper limits in the length that males can attain. That also adds uncertainty in the analyses. There were also concerns about age composition in fisheries data and as a result, the final model did not include the marginal age composition from those data.

The results of the model indicated that there was less information in the Southern input data than that in the Northern dataset. Also, the age and length data seem to be influential while likelihood profile highlighted some conflicting trends in the input data.

Pacific Ocean Perch

Although the model offers the option to estimate natural mortality and steepness, both were fixed because the input data did not provide enough information about the values of those two parameters (almost flat likelihood surface). The values used were 0.054 for natural mortality and 0.72 for steepness. The latter is considerably higher than the values found in the previous assessment which was able to estimate a value of steepness.

Overall, there is considerable uncertainty about a number of parameters in the model and the data is not informative enough to provide signals about most probable values. Given that key model parameters have been fixed, sensitivity analyses were conducted to understand the impact they had on model results. The sensitivity runs showed that the model results are sensitive to the choice of value for natural mortality and steepness. With regards to steepness, the input data support high variability in recruitment which makes the link between data-supported recruitment and what the Beverton-Holt equation predicts very loose. That makes it very difficult to calculate steepness and creates another level of uncertainty which has significant impact in model predictions as shown in the sensitivity analyses conducted. That is, the estimate of current population was almost halved when the steepness was given the value of 0.4 which is what the previous assessment used.

Model results also showed that the triennial survey supported a more pessimistic stock trend than that of the other data adding to the uncertainty characterizing the status predictions, The STAT ran analysis using only a subset of the CPUE series to depict that sensitivity of stock status results to the choice of CPUEs they used in the base

case model. Again, as with changes in steepness, the model prediction for the current status of the stock showed a decline when only the triennial survey was used.

The final base case model did not include the triennial survey as the STAT believed that this survey was not very effective in sampling the POP population.

However, concerns were also raised for the updated base case model which predicts that the population is at almost unexploited level. Although, there was a general agreement that those results seemed unrealistically optimistic, it was not possible to produce an alternative option that would be more plausible than what was already proposed.

There was also a clear tradeoff between assumptions about retention of fish and the selectivity curves the model estimated, and that also led to differences in fishery selectivity and retention curves between the current assessment and the previous one. As there is no strong evidence to define the level of retention across fisheries, and over the years it also brings some uncertainty in the analysis.

As the STAT was using the stock synthesis model, they needed to provide length and age composition data, and also used conditional age-length data to calculate growth. However, there was very little information in those data and the model results changed depending on how those data were used in the model, e.g. treating the NWFSC survey age data as conditional age-at-length data or as marginal.

The use of the SS framework meant that the team was trying to inform a model that required high level of detail to adjust all the processes of population dynamics it was simulating with data that were not detailed and informative enough to do so. Therefore, they had to fix parameters and spend a lot of time trying to find ways to make the data meet the standards of the SS. This is a useful exercise, but a simpler model might have produced results of equal value. To that effect, the STAT was asked to run a simpler model to avoid using some of the data that might not be considered so reliable. Some analyses were conducted, but it could not be fully explored due to time limitations.

Plots showing likelihood profiles of different model components also highlighted the uninformative nature of CPUE data and the influence of age composition data and associated recruitment information in driving model results.

TOR 4. Provide constructive suggestions for current improvements if technical deficiencies or major sources of uncertainty are identified.

The issues identified above were explored during the review meeting and suggestions for improvements or testing of alternative options were recommended for both species. So, in terms of immediate improvements, those would be for the STAT to capture the changes identified during the meeting (also see TOR 7). Equally, given that concerns were raised about the limited time the STAT had to incorporate some of the data into the model, it will be recommended that a more thorough consideration of the data series is done in collaboration with the people who provided the data to ensure that they have been used correctly.

For POP, there was significant variation in the recruitment predicted by the model and that mainly came from length and age composition data. It will be advisable that some additional exploration of those datasets is done to ensure that any biases or errors are addressed.

Also, it will be useful to provide a more detailed description of the dataset that was used to derive the max age, and also the age and length compositions of the species including some information to help understand the extent to which the data represent the full spatial range of the species.

After the review meeting was completed, there were still problems with an element of the Lingcod analyses (forecasting) and some mistakes were identified in one of the surveys (wrong assignment of gender for WA recreational). So, those also needed to be addressed.

TOR 5. Determine whether the science reviewed is considered to be the best scientific information available.

For both species, the assessment teams did their best to utilize the most up to date data and combine information from different resources, and their efforts are commended. The parametrization of both models changed during the meeting with some input data being dropped completely from the base case and other sources of data being corrected. The updated models reflect the data and processes in which the STAT had more confidence, and although a number of parameters are still not well defined, sensitivity analyses provided a useful perspective into the effect they have in model results. Overall, this reflects the very good effort that the teams made to produce the best assessment framework for the two species considered, and make use of the best scientific information available to construct management recommendations.

TOR 6. When possible, provide specific suggestions for future improvements in any relevant aspects of data collection and treatment, modeling approaches and technical issues, differentiating between the short-term and longer-term time frame.

This section starts with a recommendation that applies to both species and then provides species-specific suggestions.

On general issues, the most important thing to highlight is that the STAT seemed to have spent a lot of time learning the SS package and trying different parametrisations and features, but that did not seem to be the case for the input data (e.g. surveys data). That was also reflected in the decision of the STAT to change the series they used in the final model from those that had been put forward in the original assessment reports. Also, following the review meeting, further exploration of some of the input data revealed that there were not used correctly. I understand that the STAT had limited time to familiarise themselves with the data and therefore, they had to do the best within the constraints they operated. However, it is of little value employing such a sophisticated quantitative approach if the basic blocks that will underpin it are not sound and analysed/used properly. I strongly recommend that more time be spent by the STAT and other people who provide these data to understand/analyse the data and document their contribution to the assessment of the specific species and how relevant/representative they are for that species. This is in addition to a more general comment I have made under ToR 7 about documentation.

Lingcod

Short-term
The review meeting highlighted problems with non-randomly sampling fish to age, and certain datasets were identified for which that issue is of more concern (e.g. Oregon recreational fishery data). These datasets need to be revisited to adapt the analyses to account for the bias in fish sampling. Also, this highlights the need for a uniform

approach to sampling fish and for communicating the objectives and way in which sampling was done.

The Wadsworth survey was not included in the analyses because there was not enough time for the STAT to explore/understand it. Therefore, time needs to be allocated to explore this index and possible analyses that can be done.

There were also concerns about aging techniques and compatibility between labs so, it is recommended that a validation exercise is taken forward to identify any discrepancies and align procedures or characterise the level of error.

The maturity curve was calculated using only data from the recent period, and that might introduce bias in the calculations related to exploitation. It is recommended that a dataset that spans a bigger time period be used to estimate maturity and ascertain whether it has changed over the years.

Longer-term

It will be very useful to have information about the distribution of suitable habitat for Lingcod as such information could also provide an insight into the size of the Northern population relative to the Southern one. Thus, it could be an alternative way to get a crude estimate of the size of one stock relative to the other.

It is also recommended that further consideration is given to the question about stock connectivity, both in terms of data that could be collected to inform that discussion and modelling tools that could support further exploration.

Also, the model did not account for the nest-protecting behaviour that adult male fish exhibit and which might make them more vulnerable or differentiate them from the dynamics of female lingcod. It is not clear at present how important this might be or the best way to capture that process into the model, and so, further work is needed to improve understanding of this process and its impacts.

Pacific Ocean Perch

Short-term

During the meeting, the STAT noted that there were issues with catch reporting, especially relating to catches taken in California waters but landed in Oregon and which might not have been recorded properly. Further work to clarify and, if needed, adjust catch records is recommended.

Sensitivity analyses showed that the model was driven by age and length composition data. However, those data were limited, and also there was not much confidence in some of them. Given the limited information the model gets from CPUE indices, it is important that existing data are thoroughly reviewed to ensure that any bias or errors are removed. Also, I suggest that a robust length and age data collection program be put in place to improve the quality and extend of such data.

Longer-term

Although I understand that this issue has come up in the past, I will recommend that some additional work/data collection is done to check whether recruitment patterns in Canada and west coast of the USA show the same trends and provide any insight about connectivity between the two areas.

The model was predicting high recruitment variability which mainly was coming from the length/age data, but there was not a recruitment index that could verify that. Further work to understand whether such an index can be created, or explore the links to environmental and other drivers that might affect recruitment will improve the robustness of the model predictions.

TOR 7. Provide a brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations.

Lingcod

The STAT presented the data and model adopted for the assessment. That covered input data, assumptions and parametrization of the model including results. The model proposed in the assessment report sent to the reviewers was updated for both stocks to address issues identified since the assessment report was sent.

Those were:

For the North lingcod stock:

- There was a correction made to WA recreational landings (2015-2016);
- The model was set up to estimate all growth parameters except female L at Amax (female k was fixed in the model included in the assessment report);
- The model was run to estimate M and h with growth fixed.

For the South Lingcod stock:

- An early CA dockside index was added;
- There was a lambda miss-specification for conditional composition data which was fixed;
- The values of M and h were fixed at those found for the north model;
- Field CA commercial landings were added.

There was not an updated version of the report. Instead, STAT presented the new configuration and data in a PowerPoint presentation.

The results of the model for both stocks differed from those of the model sent to the reviewers, and the STAT presented the change contributed to each of the modifications they did. The Panel identified a number of changes or explorations they wanted to do, that included rerunning the VAST standardization package for each of the two West coast-wide surveys, but using the Northern data separately from the Southern. This was to align with the overall approach of the assessment which assumed complete separation of the two stocks. The results showed that the split changed the results for the triannual early for both north and south, but only for the

South for the triennial Late index. The other 3 CPUEs remained almost unchanged. However, there was little change in the overall trend for all 6 indices.

The Panel pointed out that the model description in relation to selectivity was not clear and requested additional work to fill the gaps. That included a description and an explanation of the selectivity blocks used in the model. The Panel also requested runs with the recruitment deviations incorporated from the beginning of the time series and incorporation of catches in the early years of the calculations for the Southern stock, for which catch reconstruction extend back to only 1930s. This was mainly to check if there is information in the data to inform the scale of deviations in early years and create a more realistic scenario for the Southern stock as the general view was that the stock was exploited pre-1930s.

A number of requests were made to understand the length and age composition of data collected as there were concerns that sampling of fish for aging might not have been random. Further exploration revealed that the sampling of commercial fisheries was not random, and in some cases, younger fish were given preferences when fish were chosen for aging. Therefore, it was considered that fishery-dependent age composition data should be excluded from the analyses.

During discussions, the STAT indicated that they had more confidence in the data that came from the NWFSC Trawl Survey; as such and to test the impact of each data source, the STAT was asked to run the assessment for both Northern and Southern stocks using only the NWFSC index and then add other indices one at a time.

The Panel also questioned how dressed fish are captured/handled in the data and it was confirmed that they had been excluded from the data (about 1% of the data). It was indicated that it is probable that a greater proportion of older than younger fish fall in that category, and so, their exclusion could have excluded older fish more than younger ones.

Consideration of the southern model's results revealed that the CPUE from hook and line was in biomass when it should have been in numbers. The index was reanalyzed and the updated one in numbers was used for the calculations. The updated index did not change the overall results of the model.

Other things that were clarified during the review were about differences in the spatial distribution the surveys considered in the Southern model, and also in relation to one more CPUE index which was available (Wadsworth) but was not used as the STAT did not have time to explore it.

Following a series of sensitivity and exploration runs, the STAT transitioned to updated base case models for both stocks that were different from the original ones. Changes included exclusion of some age composition data and fixing the values of key biological parameters such as natural mortality and steepness. The Panel also discussed with the STAT the sensitivity analyses to conduct to represent uncertainty in forecasting and management advice. That discussion continued and after the review meeting, it was decided the standard deviation on 2017 spawning biomass would be used to define the range of values to consider.

Pacific Ocean Perch

At the beginning of the meeting the STAT indicated that the model was updated to address some issues identified since the assessment report was sent. Those were:

- Addition to California historical landings 1948-1968: corrections totaling 10 mt;
- Survey catch removal correction: Stock Synthesis was not removing catches for survey fleets; and
- Weight-at-length: Small correction to the weight-at-length values for females and males.

There was not an updated version of the report. Instead, STAT presented the new configuration and data in a PowerPoint presentation.

The results of the updated model were very similar to the ones sent to the reviewers. The STAT team gave presentations describing the data and model on the first day and the Panel formulated requests for further analyses. Three sets of requests were formed and responses were presented throughout the week. A summary of the issues covered is provided below.

Given that many of the model inputs had been changed, the Panel asked for continuity analysis to see how the previous model (one used in 2011) would behave with the new data. There were also a number of requests for additional calculations to help understand the impacts that different inputs to the model had on its results. There were questions about the NWFSC series, about how it was constructed (age composition) and how it affected the model. On the latter, the STAT was asked to run the model with the NWFSC index only, and also with the NWFSC and the triannual survey index to check how different the final results would be from the model presented at the beginning of the review meeting if only fishery-independent CPUE indices were used.

The STAT were also requested to adjust SS so that it behaves like a standard age-structured production model to test whether the trends would change when the model is not fit to age composition data. A number of other sensitivity runs following the same logic (test the impact that a single source of data had in the model) were also requested, and individual likelihood components coming from different model parametrizations were examined to understand the impacts of each component and look for conflicting trends supported by different data sources. On the latter, the triennial survey appeared to support trends that are contradictory to those supported by other indices and composition data; therefore, the model was run using the triannual survey as CPUE index to understand the status of the stock that it supported. Selectivity and retention assumptions for the fishery were also explored and compared to those used in 2011. The assumptions for retention differed between 2011 and 2017, and that also supported different selectivity at age for the 2017 model which had slightly smaller selectivity for younger fish (length smaller than 25 mm) than the 2011 model.
As mortality and steepness were main sources of uncertainty, the model was run for different values of those parameters to reflect the uncertainty they bring in model results. Using blocking to represent changes in selectivity over the years was also suggested as the Panel considered it to reflect the changes in regulations over the years.

This process led to the STAT suggesting a new base model which included all the corrections identified at the beginning of the review meeting, excluded the triennial surveys and used blocking to describe changes in selectivity of the fishing fleet over the years. The Panel agreed with the selection, and discussed how to define the envelope of uncertainty to capture the uncertainty in model results. Both natural mortality and steepness were considered appropriate to define the envelope, and the range of values were defined by the 12.5 and 87.5 quantiles of the priors from steepness and natural mortality.

General recommendations

As I was the constant reviewer some of my comments below apply to this assessment review, but also represent my overall impression of processes and issues across assessments reviewed.

Several surveys were presented and used for the stock assessments and have been collecting data for many years. Although a short description of each of them was included in the assessment report, there is the need for a more in-depth description of those surveys so, their scope, spatial distribution, main characteristics and type of fish for which it is appropriate can be better understood. At present, the descriptions available are fragmented, and it is even more difficult to understand how those surveys perform individually but also cumulatively with regards to species assessed. Therefore, it is recommended that a standalone document providing details about each survey be prepared, and that serves as the basis for any further, maybe species-specific, description of the surveys that is included in the assessment report.

The assessment teams continued their analyses even after they had submitted the draft assessment to the STAR Panel. As such, by the time the Panel meeting took place, there were new results and modifications of the model. However, a revised assessment report to capture the updated model or results was not provided prior to the meeting. I understand that the assessment team is not allowed in principle to modify the assessment report once it has been sent to the reviewers (but I believe this is because the report should capture the final version of the model). This means that the version of the analyses the reviewers consider is superseded by a different one by the time the meeting starts creating confusion and makes the whole process less efficient. So, I strongly recommend that if assessment teams are allowed to conduct further analyses and present new results on the first day of the meeting, then they should also produce a revised report. They should also notify the reviewers if they are not very confident in their analyses and plan to consider further improvements in the period between the time the report is sent to the reviewers and the Panel meeting takes place. This of course does not address the issue highlighted during the meeting which was that the STAT did not have enough time to do the assessments and should also be addressed.

Also, it would be very useful if all responses to Panel's requests are provided with some text explaining what was done and what the results are, what changes to the results are made, etc., instead of just providing a figure or a table without a written explanation. The Panel is expected to provide its requests to the assessment team in writing and provide some rationale for the request. It will be useful if the responses were also structured in a similar way, i.e. the results and what it means with regards to what the request wanted to achieve/clarify.

Conclusions/Recommendations

The benchmark stock assessments of lingcod and Pacific ocean perch (POP) review meeting took place in Seattle between June 26th and June 30th, 2017. The focus of the review was the assessment work done for the two stocks of lingcod and one stock of POP. Both species were assessed using a length- and age-structured modelling software called Stock Synthesis, which also allows for age and length composition data to be incorporated into the analyses.

A diverse set of data were used as input to the models covering fishery-dependent and fishery independent sources. However, the data were not very informative and that meant that a number of assumptions had to be made about key processes and parameters to get the models to converge. Notwithstanding that, the STAT made a commendable effort to utilise all data available and maximise the knowledge that could be derived from them.

The analyses predicted that the stock of POP species was in a better shape than the previous assessment predicted and projections indicated that the stock could support higher catches than those currently taken. For lingcod, the results were more pessimistic than those from the previous assessment, indicating that one of the stocks has not recovered yet. Some concerns still remain about the realism in the model predictions and issues with input data, and therefore, sensitivity analyses were employed to address that as much as possible. Overall, the results provide a useful picture of the uncertainties in the knowledge about these species and predicted trends in their stocks and could inform decision making. However, the review also identified a number of areas that, if improved, would strengthen the quality and robustness of the assessment results.

A list of the recommendations made under each of the ToR above are summarised here starting with general ones and then listing species specific ones.

Recommendation 1: For both species, it will be useful to provide a more detailed description of the datasets used to derive the max age, and also the age and length compositions of the species including some information to help understand the extent to which the data represent the full spatial range of the species.

Recommendation 2: I strongly recommend that more time be spent by the STAT and other people who provide these data to understand the data and document their contribution to the assessment of the specific species and how relevant/representative they are for that species. This is in addition to a more general comment I have made about documentation below.

Recommendation 3: It is recommended that existing age and length composition data be thoroughly reviewed to ensure that any bias or errors are removed. Also, I suggest that a robust length and age data collection program be put in place to improve the quality and extend of such data.

Recommendation 4: It is recommended that a standalone document providing details about each survey should be prepared, and that serves as the basis for any further, maybe species-specific, description of surveys included in the assessment report.

Recommendation 5: I strongly recommend that if assessment teams are allowed to conduct further analyses and present new results on the first day of the meeting, then they should also produce a revised report and notify the reviewers if they are not very confident in their analyses and plan to consider further improvements in the period between the time the report is sent to reviewers and the Panel meeting takes place.

Recommendation 6: It would be very useful if all responses to Panel's requests are provided with some text explaining what was done and what the results are, what changes to the results were made, etc., instead of just providing a figure or a table without a written explanation.

Recommendation 7: Lingcod: The review meeting highlighted problems with non-randomly sampling fish to age and certain datasets were identified for which that issue is of more concern (e.g. Oregon recreational fishery data). These datasets need to be revisited to adapt the analyses to account for the bias in fish sampling. Also, this highlights the need for a uniform approach to sampling fish and for communicating the objectives and way in which sampling was done.

Recommendation 8: Lingcod: Time needs to be allocated to explore the Wadsworth survey index and possible analyses that can be done to ascertain whether it could be used in the next assessment.

Recommendation 9: Lingcod: It is recommended that a validation exercise for the aging process be taken forward to identify any discrepancies and align procedures or characterise the level of error between labs.

Recommendation 10: Lingcod: It is recommended that a dataset that spans a bigger time period be used to estimate maturity and consider whether it has changed over the years.

Recommendation 11: Lingcod: It will be very useful to have information about the distribution of suitable habitat for Lingcod, as such information could also provide an insight into the size of the Northern population relative to the Southern one.

Recommendation 12: Lingcod: It is recommended that further consideration be given to the question about stock connectivity both in terms of data that could be collected to inform that discussion, and modelling tools that could support further exploration.

Recommendation 13: Lingcod: Further work is needed to improve understanding of the nest-protecting behaviour exhibited by adult males and its impacts.

Recommendation 14: POP: It was noted that there were issues with catch reporting, especially relating to catches taken in California waters but landed in Oregon and which might not have been recorded properly. Further work to clarify, and, if needed, adjust catch records is recommended.

Recommendation 15: POP: There was significant variation in the recruitment predicted by the model and that mainly came from length and age composition data. It is recommended that some additional exploration of those datasets be done to ensure that any biases or errors are addressed.

Recommendation 16: POP: I will recommend that some additional work/data collection is done to check whether recruitment patterns in Canada and west coast of the USA show the same trends and provide any insight about connectivity between the two areas.

Recommendation 17: POP: Further work to understand whether a recruitment index could be created or explore the links to environmental and other drivers that might affect recruitment could improve the robustness of the model predictions.

Recommendation 18: POP: It is worth monitoring the catches of the shrimp fishery and take their impact into account especially when a large new POP cohort has been detected.

Appendix 1: Bibliography

Anon. 2017. Report of the Groundfish Historical Catch Reconstruction Workshop. Pacific Fishery Management Council. Portland, OR. 1-3 November 2016.

Anon. 2017. Report of the Groundfish Productivity Workshop of the Pacific Fishery Management Council's Scientific and Statistical Committee. NOAA Fisheries, Alaska Fisheries Science Center, Seattle, Washington. December 6-8, 2016.

Anon. 2017. Scientific and Statistical Committee Report on Reports and Recommendations from Groundfish Science Workshops and Methodology Reviews.

Anon. 2017. Summary Minutes. 2017 Groundfish Pre-Assessment Workshop. March 21-22, 2017. Pacific Fishery Management Council.

Anon. No date. Accepted Practices Guidelines for Groundfish Stock Assessments.

Anon. No date. Summary table of Triennial Shelf Survey and NWFSC Shelf/Slope Survey. This table was completed with input from Victor Simon and Aimee Keller (NWFSC) and Mark Wilkins (AFSC).

Dorn. M. 2016. Design considerations for PFMC groundfish stock assessments and reference points. Powerpoint presentation. Groundfish Productivity Workshop. Seattle, Washington.

Hamel, O. S. 2015. A method for calculating a meta-analytical prior for the natural mortality rate using multiple life history correlates. *ICES Journal of Marine Science*, **72**: 62–69.

Hamel, O. S. 2017. An update to developing a prior for the natural mortality rate for fishes for use in the 2017 assessment cycle for the U.S. West Coast. Unpublished document.

Karnowski, M., Gertseva, V., and Stephens, A. 2014. Historical Reconstruction of Oregon's Commercial Fisheries Landings. Fish Division, Oregon Department of Fish and Wildlife. Information Reports Number 2014-02. 52 pp.

Methot Jr, R. D. and Wetzel, C. R. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research, 142: 86– 99.

Methot Jr, R. D. and Wetzel, C. R. 2013. Appendix A: Technical Description of the Stock Synthesis assessment program. Supplementary data. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research, 142: 86– 99.

Methot, R. D., A'mar, T., Wetzel, C., and Taylor, I. 2017. Stock Synthesis. User Manual. Version 3.30.04. June 1, 2017. NOAA Fisheries, Seattle, WA USA.

NOAA. No date. Information Quality Guidelines.

Pacific Fishery Management Council. 2016. Terms of Reference for the Groundfish and Coastal Pelagic Species Stock Assessment Review Process for 2017-2018.

Thorson, J. T. 2017. Estimating a Bayesian prior for steepness in Pacific rockfishes (*Sebastes* spp.) 1 off the U.S. West Coast for the 2017 assessment cycle. Presented to SSC Assessment Methods Review. Jan. 11, 2017.

Thorson, J. T., and Barnett, L. A. K. Comparing estimates of abundance trends and distribution shifts using single- and multispecies models of fishes and biogenic habitat. – ICES Journal of Marine Science, doi:10.1093/icesjms/fsw193.

Thorson, J. T., Shelton, A. O., Ward, E. J., and Skaug, H. J. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. – ICES Journal of Marine Science, doi: 10.1093/icesjms/fsu243.

Thorson, J. T., Dorn, M. W., and Hamel, O. S. No date. Steepness for West Coast rockfishes: Results from a twelve-year 1 experiment in iterative regional meta-analysis. Draft: For Review During PFMC Star Panel Reviews In 2017.

Background – STAR Panel 1

Anon., No date. Summary of California Recreational Management Measures Affecting Lingcod.

Anon. No date. [Federal] Regulations [affecting Lingcod] by Multiple Parameters.

Anon. No date. Excel workbook, Lingcod_SupportingMaterials_WA_Lingcod-RecRegsArea1-4.xlsx. [Annual regulations for recreational fishing for Lingcod].

Chen, Y. 2011. CIE Independent Peer Review Report on Pacific Ocean Perch and Petrale Sole STAR Review. Report to CIE.

Hamel, O. S., Sethi, S. A., and Wadsworth, T. F. 2009. Status and Future Prospects for Lingcod in Waters off Washington, Oregon, and California as Assessed in 2009. Lingcod Assessment 2009 – Final – SAFE Version.

Hamel, O. S. 2009. Status and Future Prospects for the Pacific Ocean Perch Resource in Waters off Washington and Oregon as Assessed in 2009. Pacific Ocean Perch Assessment 2009 – Final – SAFE Version.

Hamel, O. S., and Ono, K. 2011. Stock Assessment of Pacific Ocean Perch in Waters off of the U.S. West Coast in 2011.

Jagielo, T. H. and Wallace, F. R. 2005. Assessment of Lingcod (*Ophiodon elongatus*) for the Pacific Fishery Management Council in 2005.

Maguire, J.-J. 2009. Report on the cabezon and lingcod Stock Assessment Review (STAR) Panel, July 27 - 31, 2009, Seattle, WA. Report to CIE.

Smith, S. 2009. Report for the Center of Independent Experts on the Stock Assessment Review (STAR) Panel for Cabezon and Lingcod (July 27 to 31, 2009). Report to CIE.

SSC. 2009. Scientific and Statistical Committee Report on Stock Assessments for 2011-2012 Groundfish Fisheries. Supplemental SSC Report, June 2009.

STAR Panel. 2005. STAR Panel Report. Lingcod. Northwest Fisheries Science Center, Seattle, Washington, USA. 15-19 August 2005.

STAR Panel. 2005. Lingcod. STAR Panel Report. Alaska Fisheries Science Center, Seattle, Washington, USA. September 26-30, 2005.

STAR Panel. 2009. Lingcod. STAR Panel Report. Deca Hotel, Seattle, Washington. July 27-30, 2009.

STAR Panel. Pacific Ocean Perch. Stock Assessment Review (STAR) Panel Report. Hotel Deca, Seattle, Washington. 20-24 June 2011.

Stokes, K. 2011. Report on the Stock Assessment Review (STAR) Panel for Pacific Ocean Perch and Petrale Sole. Report to CIE.

Whitman, A. 2017. Summary: Sport Regulation Changes relevant to Lingcod, and other items that may be relevant to CPUE development.

Draft Lingcod Assessment

4_North_v7-hess-ToSTAR. Folder containing input files for SS3 analysis of northern stock of lingcod and associated output tables and figures.

4_North_w2.1-hess-AltToSTAR.

5_South_e_base-hess-ToSTAR. Folder containing input files for SS3 analysis of southern stock of lingcod and associated output tables and figures.

5-South_f5-hess-AltToSTAR. Folder containing alternative input files.

Haltuch, M. A., Wallace, J., Akselrud, C. A., Nowlis, J., Barnett, L. A. K., Valero, J. L., Tsou, T.-S., and Lam, L. 2017. Draft 2017 Lingcod Stock Assessment, with associated figures and tables.

Lingcod all lengths prop vs. lengths with ages prop
Model parameters_STAR
NWFSC Survey_North Lingcod Catch
NWFSC Survey South Lingcod Catch

Future Research Recommendations for the Lingcod Assessment

Draft Pacific Ocean Perch Assessment

CSV files, POP_natage_f.csv and POP_natage_m.csv, containing number-at-age data for Pacific Ocean Perch.

SS3 input files for Pacific Ocean Perch stock assessment.

Recommendations for Future Research for the POP Assessment.

Wetzel, C. R. and Cronin-Fine, L. 2017. Status of Pacific ocean perch (*Sebastes alutus*) along the US west coast in 2017.

Appendix 2. Statement of Work for Dr Panagiota Apostolaki

External Independent Peer Review by the Center for Independent Experts

Stock Assessment Review (STAR) Panel 1

Background

The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NMFS science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards.
(http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf).
Further information on the CIE program may be obtained from www.ciereviews.org.

Project Description

The National Marine Fisheries Service and the Pacific Fishery Management Council will hold stock assessment review (STAR) panels in 2017 to evaluate and review benchmark assessments of Pacific coast groundfish stocks. The goals and objectives of the groundfish STAR process are to:

1) ensure that stock assessments represent the best available scientific information and facilitate the use of this information by the Council to adopt OFLs, ABCs, ACLs, (HGs), and ACTs;
2) meet the mandates of the Magnuson-Stevens Fisheries Conservation and Management Act (MSA) and other legal requirements;
3) follow a detailed calendar and fulfill explicit responsibilities for all participants to produce required reports and outcomes;
4) provide an independent external review of stock assessments;
5) increase understanding and acceptance of stock assessments and peer reviews by all members of the Council family;
6) identify research needed to improve assessments, reviews, and fishery management in the future; and
7) use assessment and review resources effectively and efficiently.

Benchmark stock assessments will be conducted and reviewed for lingcod and Pacific ocean perch. Lingcod has been an important groundfish target species along the west coast of the United States, ranking in the top-6 of importance for commercial, recreational, and tribal fisheries. This will be the first assessment for lingcod since a

benchmark assessment was completed in 2009. During the last benchmark assessment, the sensitivity results showed high uncertainty in age data as well as in the status of the population in southern California. A substantial effort is currently underway to age lingcod to reduce uncertainties relating to age and growth data and improve recruitment estimates.

Pacific ocean perch has been managed under a rebuilding plan for over a decade and, while not expected to be rebuilt for several more decades (2051), was identified as a strong candidate for assessment during the Pacific coast groundfish regional stock assessment prioritization process, which was based on the national stock assessment prioritization framework (http://www.st.nmfs.noaa.gov/Assets/stock/documents/PrioritizingFishStockAssessments_FinalWeb.pdf. Pacific ocean perch was assessed as a benchmark assessment in 2011 and a catch-only rebuilding projection update in 2015 to monitor the rebuilding progress and provide updated scientific-based advice for management. This stock has recently become more constraining on elements of the Pacific hake fishery.

Assessments for these two stocks will provide the basis for the management of the groundfish fisheries off the West Coast of the U.S. including providing scientific basis for setting OFLs and ABCs as mandated by the Magnuson-Stevens Act. The technical review will take place during a formal, public, multiple-day meeting of fishery stock assessment experts. Participation of external, independent reviewer is an essential part of the review process. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**.

Requirements for CIE Reviewers
NMFS requires two CIE reviewers to participate in this stock assessment review panel. One CIE reviewer shall conduct an impartial and independent peer review of the two assessments described above and in accordance with the SoW and ToRs herein. Additionally, a second "consistent" CIE reviewer will participate in all STAR panels held in 2017 and the SOW and ToRs for the "consistent" CIE reviewer are included in a separate SoW (See **Attachment A**).

Both CIE reviewers shall be active and engaged participants throughout panel discussions and able to voice concerns, suggestions, and improvements while respectfully interacting with other review panel members, advisors, and stock assessment technical teams. The CIE reviewers shall have excellent communication skills in addition to working knowledge and recent experience in fish population dynamics, with experience in the integrated analysis modeling approach, using age- and size-structured models, use of *Markov Chain Monte Carlo (*MCMC) to develop confidence intervals, and use of Generalized Linear Models in stock assessment models.

Statement of Tasks
The CIE reviewers shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

Pre-review Background Documents: At least two weeks before the peer review, the contractor will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. The CIE reviewer shall read all documents in preparation for the peer review.

Documents to be provided to the CIE reviewers prior to the STAR Panel meeting include:

- The current draft stock assessment reports;
- The Pacific Fishery Management Council's Scientific and Statistical Committee's Terms of Reference for Stock Assessments and STAR Panel Reviews;
- Stock Synthesis (SS) Documentation
- Additional supporting documents as available (including previous stock assessments and STAR panel reports).
- An electronic copy of the data, the parameters, and the model used for the assessments (if requested by reviewer).

Panel Review Meeting: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein. Each CIE reviewer shall actively participate in a professional and respectful manner as a member of the meeting review panel, and their peer review tasks shall be focused on the ToRs as specified herein.

Contract Deliverables - Independent CIE Peer Review Reports: The CIE reviewers shall complete an independent peer review report in accordance with the SoW. Each CIE reviewer shall complete the independent peer review according to required format and content as described in **Annex 1**. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in **Annex 2**.

Other Tasks – Contribution to Summary Report: The CIE reviewers may assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review. The CIE reviewers are not required to reach a consensus, and should provide a brief summary of each reviewer's views on the summary of findings and conclusions reached by the review panel in accordance with the ToRs.

Timeline for CIE Reviewers
The following chronological list of tasks shall be completed by each CIE reviewer in a timely manner as specified in the **Schedule of Milestones and Deliverables**.

4) Conduct necessary pre-review preparations, including the review of background material and reports provided in advance of the peer review.
5) Participate during the STAR Panel 1 review meeting in **scheduled in Seattle, Washington during the dates of June 26-30, 2017** as specified herein, and conduct an independent peer review in accordance with the ToRs (**Annex 2**).
6) No later than **July 14, 2017**, each CIE reviewer shall submit their draft independent peer review report to the contractor. Each CIE report shall be written using the format and content requirements specified in **Annex 1**, and address each ToR in **Annex 2**.

Foreign National Security Clearance
When reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for reviewers who are non-US citizens. For this reason, the reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall

be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/ and http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration-system.html. The contractor is required to use all appropriate methods to safeguard Personally Identifiable Information (PII).

Place of Performance
The CIE reviewers shall conduct an independent peer review during the panel review meeting scheduled in **Seattle, Washington during the dates of June 26-30, 2017**.

Period of Performance
The period of performance shall be from the time of the award through August 18, 2017. Each reviewer's duties shall not exceed 14 days to complete all required tasks.

Schedule of Milestones and Deliverables
The contractor shall complete the tasks and deliverables described in this SoW in accordance with the following schedule.

May 22, 2017	Contractor selects and confirms reviewers
June 12, 2017	Contractor provides the pre-review documents to the reviewers
June 26-30, 2017	Each reviewer participates and conducts an independent peer review during the panel review meeting
July 14, 2017	Contractor receives draft reports
July 31, 2017	Contractor submits final reports to the Government

Applicable Performance Standards
The acceptance of the contract deliverables shall be based on three performance standards: (1) The reports shall be completed in accordance with the required formatting and content in **Annex 1**; (2) The reports shall address each ToR as specified **Annex 2**; and (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel
All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (http://www.gsa.gov/portal/content/104790). International travel is authorized for this contract. Travel is not to exceed $8,200.

Restricted or Limited Use of Data
The contractors may be required to sign and adhere to a non-disclosure agreement.
NMFS Project Contacts:
Stacey Miller, NMFS Project Contact
National Marine Fisheries Service,
2032 SE OSU Drive
Newport, OR 97365
Stacey.Miller@noaa.gov

Phone: 541-867-0535

Jim Hastie
National Marine Fisheries Service,
2725 Montlake Blvd. E,
Seattle WA 98112
Jim.Hastie@noaa.gov
Phone: 206-860-3412

Annex 1: Format and Contents of CIE Independent Peer Review Report

1. The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

   a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including providing a brief summary of findings, of the science, conclusions, and recommendations.

   b. Reviewers should discuss their independent views on each ToR even if these were consistent with those of other panelists, and especially where there were divergent views.

   c. Reviewers should elaborate on any points raised in the Summary Report that they feel might require further clarification.

   d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

   e. The CIE independent report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The CIE independent report shall be an independent peer review of each ToRs, and shall not simply repeat the contents of the summary report.

3. The reviewer report shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of the CIE Statement of Work
   Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

Annex 2: Terms of Reference for the Peer Review

Stock Assessment Review (STAR) Panel 1

1. Become familiar with the draft stock assessment documents, data inputs, and analytical models along with other pertinent information (e.g. previous assessments and STAR panel report when available) prior to review panel meeting.
2. Discuss the technical merits and deficiencies of the input data and analytical methods during the open review panel meeting.
3. Evaluate model assumptions, estimates, and major sources of uncertainty.
4. Provide constructive suggestions for current improvements if technical deficiencies or major sources of uncertainty are identified.
5. Determine whether the science reviewed is considered to be the best scientific information available.
6. When possible, provide specific suggestions for future improvements in any relevant aspects of data collection and treatment, modeling approaches and technical issues, differentiating between the short-term and longer-term time frame.
7. Provide a brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations.

Annex 3: Tentative Agenda

TBD

Stock Assessment Review (STAR) Panel 1
NMFS Northwest Fisheries Science Center
2725 Montlake Blvd NE
Seattle, WA 98112

June 26-30, 2017

Appendix 3: STAR Panel Membership
In alphabetical order

Dr. Panayiota Apostolaki	Center for Independent Experts
Dr. Norman Hall	Center for Independent Experts
Dr. Kevin Piner	NOAA Fisheries, Southwest Fisheries Science Center, SSC
Dr. David Sampson	Oregon State University, Scientific and Statistical Committee (SSC), STAR Panel Chair